Using the LLNA to Categorize Strong Skin Sensitizers

J Strickland[1], J Matheson[2*], M Wind[2*], A Jacobs[3], F Stack[1], D Allen[1], W Stokes[4]

[1]ILS, Inc., RTP, NC; [2]US CPSC, Bethesda, MD; [3]US FDA, Silver Spring, MD; [4]NICEATM/NIEHS/NIH/DHHS, RTP, NC

Abstract[1]

According to the U.S. Bureau of Labor Statistics, allergic contact dermatitis (ACD) is one of the most common types of occupational disease. Because the prognosis of ACD is poor, prevention is imperative. Criteria have recently been adopted to distinguish strong sensitizers from other sensitizers based on human, guinea pig, and the murine local lymph node assay (LLNA) data. Substances with positive responses in the human maximization test (HMT) or human repeat insult patch test (HRIPT) at induction thresholds ≤500 μg/cm^2 are classified as strong sensitizers. Similarly, LLNA EC3 values ≤ 2% are proposed to categorize substances as strong sensitizers and LLNA EC3 values >2% to categorize substances as "other sensitizers". In order to evaluate the accuracy of the LLNA for identifying strong sensitizers as defined by human data, NICEATM and ICCVAM used a database of 112 substances with both LLNA and human data to calculate human potency classification categories (strong vs. other than strong) predicted by various EC3 values. Classifications based on EC3 values were compared to those defined by several different threshold values derived from HMT and HRIPT studies. Based on the available database, 64% of strong human sensitizers were correctly predicted using LLNA EC3 ≤ 2%, while the remaining 36% of strong sensitizers were underclassified as "other sensitizers." The current database indicates that over 1/3 of strong human sensitizers would be underclassified as weaker skin sensitizers if the LLNA is used to determine potency categories. Therefore, the LLNA should not be considered as a stand-alone test to predict skin sensitization potency. While the LLNA EC3 < 2% can be used to categorize a substance as a strong sensitizer, EC3 values greater than 2% should not be used to categorize substances as not being strong human sensitizers due to the high rate of under prediction of strong human sensitizers. Other types of supporting information (e.g., QSARs, peptide reactivity, human evidence, validated *in vitro* assays, historical data from related substances, other animal studies, etc.) should be investigated for their usefulness in increasing the accuracy of categorization criteria for strong sensitizers. Information found to be useful should be incorporated into an integrated decision strategy for categorization.

[1]The abstract has been modified slightly from the version submitted.

Introduction

- Allergic contact dermatitis (ACD) is one of the most common types of occupational disease. Because the prognosis is poor, prevention is imperative.
  - Prevention requires limiting human exposure to substances that are classified as potential skin sensitizers.
- The United Nations Globally Harmonized System for Classification and Labelling of Chemicals (GHS) includes criteria for classifying substances as skin sensitizers (which produce ACD) or unclassified substances (i.e., nonsensitizers) based on human and/or animal data (UN 2009).
- The GHS was revised in 2009 to include the option of further subdividing potential skin sensitizers into "strong" (1A) and "other" (1B) categories (**Table 1**).
  - Classification criteria are based on:
    - Induction concentrations in the human repeat insult patch test (HRIPT) and the human maximization test (HMT)
    - Responses in the guinea pig maximization test (GPMT) or the Buehler test (BT)
    - LLNA EC3 values (estimated substance concentration that produces a stimulation index of 3)
- This analysis examines the accuracy of the LLNA EC3 for predicting the strong and other human skin sensitizer categories based on the HRIPT or HMT induction threshold of 500 μg/cm^2 (UN 2009).

Table 1. GHS Classification Categories for Skin Sensitizers

Category	Classification Criteria	LLNA EC3	Human Evidence (HRIPT or HMT)	GPMT Response	BT Response
1: Skin sensitizer	Evidence that skin sensitization occurs in a substantial number of people, or positive results from an appropriate animal test	NA	NA	NA	NA
1A: Strong skin sensitizer	High frequency of occurrence in humans, and/or high potency in animals. May consider severity.	≤2%	Positive[1] response at ≤500 mg/cm^2	≥30% responders at ≤0.1% intradermal induction dose or ≥60% responders at >0.1% to ≤1% intradermal induction dose	≥15% responders at ≤0.2% topical induction dose or ≥60% responders at >0.2% to ≤20% topical induction dose
1B: Other skin sensitizer	Low to moderate frequency of occurrence in humans, and/or low to moderate potency in animals. May consider severity.	>2%	Positive[2] response at >500 mg/cm^2	≥30% to <60% responders at >0.1% to ≤1% intradermal induction dose or ≥30% responders at >1% intradermal induction dose	≥15% to < 60% responders at >0.2% to ≤20% topical induction dose or ≥15% at >20% topical induction dose

Abbreviations: BT = Buehler test; CPSC = U.S. Consumer Product Safety Commission; GPMT = guinea pig maximization test; HMT = human maximization test; HRIPT = human repeat insult patch test; LLNA EC3 = estimated substance concentration that produces a stimulation index of 3 in the murine local lymph node assay; NA = not applicable.

[1]Human evidence can also include diagnostic patch test data where there is a relatively high and substantial incidence of reactions in a defined population in relation to relatively low exposure or other epidemiology evidence where there is a relatively high and substantial incidence of allergic contact dermatitis in relation to relatively low exposure.

[2]Human evidence can also include diagnostic patch test data where there is a relatively low but substantial incidence of reactions in a defined population in relation to relatively high exposure or other epidemiology evidence where there is a relatively low but substantial incidence of allergic contact dermatitis in relation to relatively high exposure.

Methods

Human Test Methods

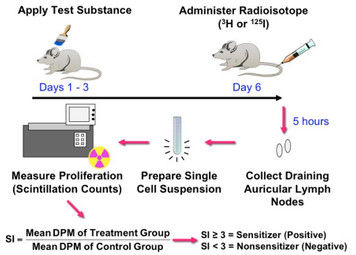
- The HMT and HRIPT tests involve the administration of occluded patches, loaded with test substance, to the skin for 5 to 9 on-and-off periods of 24-48 hours in order to attempt to induce an allergic reaction (Kligman and Epstein 1975; Politano and Api 2007).
- Following a rest period of several days, volunteers are again exposed to the test substance in an occluded patch on naïve skin for 24-48 hours.
- Skin reactions noted after patch removal suggest skin sensitization and are noted as positive reactions.
- For substances that produce no skin irritation, the HMT includes a patch pre-treatment of the skin with 5% sodium lauryl sulfate for the 24 hour period prior to the induction patch treatments in order to compromise the stratum corneum barrier (Kligman and Epstein 1975). This concentration produces a brisk dermatitis in most Caucasians.
- Induction thresholds for positive reactions are reported as micrograms of applied substance per cm^2 area of skin.

Figure 1. Collage of photographs showing a patch test (top center) surrounded by other images of dermatitis typical of ACD

LLNA Test Method

- ICCVAM evaluated the LLNA test method (see **Figure 2**) and compared the accuracy and reliability of the LLNA to guinea pig skin sensitization tests and to human data (ICCVAM 1999; Dean et al. 2001; Haneke et al. 2001; Sailstad et al. 2001). The ICCVAM evaluation concluded that:
  - The LLNA was a valid alternative to guinea pig test methods for many testing situations;
  - The LLNA reduced the number of animals required for testing while also eliminating animal pain and distress.

Figure 2. LLNA Test Method



Results

Chemical Database for Analysis

Data were obtained from published reports or data submitted to NICEATM in response to a *Federal Register* (FR) notice (72 FR 27815).

- The database included 112 substances with both LLNA and human data (ICCVAM 2008)
- The EC3 values or human thresholds for substances with multiple values were used to calculate a geometric mean* so that one LLNA EC3 and one human threshold value represented each substance.
  - Human thresholds were lowest-observed-effect levels or doses per unit area that produced a 5% response (DSA05) in the population tested.
  - Geometric means for the LLNA EC3 values were calculated using the results for the most prevalent vehicle when tests with multiple vehicles were available.
  - EC3 values ranged from 0.0028 to 88.5%; human induction threshold values ranged from 1.7 to 68966 μg/cm^2.

The 112 substances included:

- 25 strong human sensitizers (HMT or HRIPT induction threshold ≤ 500 μg/cm^2)
  - 24 LLNA sensitizers
  - 1 LLNA false negative
- 43 other human sensitizers (HMT or HRIPT induction threshold > 500 μg/cm^2)
  - 37 LLNA sensitizers
  - 6 LLNA false negatives
- 44 human nonsensitizers
  - 19 concordant LLNA negatives
  - 25 LLNA false positives (24 with EC3 >2%, 1 with EC3 ≤ 2%)

*A geometric mean is the nth root product of n numbers. For the data set [a1, a2, ..., an], it is defined by the equation:

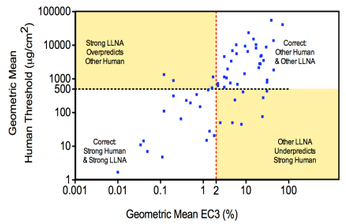
$$\left(\prod_{i=1}^{n} a_i\right)^{1/n} = \sqrt[n]{a_1 a_2 \cdots a_n}.$$

Relative Potency

61 of the 68 human sensitizers were also LLNA sensitizers, and these substances were analyzed for relative potency based on GHS potency categorization as shown in **Figure 3**.

- Includes LLNA sensitizers from the following human categories:
  - 24 strong human sensitizers
  - 37 other human sensitizers
- Excludes 7 LLNA false negatives (i.e., substances lacking EC3 values):
  - 1 strong human sensitizer
  - 6 other human sensitizers

Figure 3. Relative Potency of 61 LLNA and Human Sensitizers



Abbreviations: EC3 = estimated substance concentration that produces a stimulation index of 3 in the murine local lymph node assay; LLNA = murine local lymph node assay.

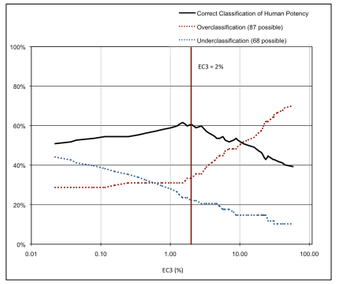
Note: The graph does not show 7 LLNA false negatives, 25 LLNA false positives, or 19 concordant LLNA negatives.

- **Figure 3** shows the geometric mean human threshold (i.e., induction concentration that produces a positive response in the HMT or HRIPT) and LLNA EC3 values for 61 LLNA and human sensitizers.
  - Human thresholds were lowest-observed-effect levels or doses per unit area that produced a 5% response (DSA05).

Potency Prediction

- To determine the ability of the LLNA EC3 to predict the human potency categories (i.e., strong or other), counts of substances above and below various EC3 cutoff values were used to calculate the overall rate of correct classification, overclassification, and underclassification. In addition, the rates of correct classification, overclassification, and underclassification for the LLNA EC3 of 2% were calculated for strong human sensitizers, other human sensitizers, and human nonsensitizers.
- **Figure 4** shows the overall rate of correct classification (combined for strong, other, and nonsensitizers for all 112 substances), overclassification (87 substances for both other and nonsensitizers), and underclassification (68 substances for both strong and other sensitizers) by the LLNA EC3.
  - The correct classification rate is maximized at EC3 values of approximately 1.5 to 2%.
  - As the LLNA EC3 increases, the underclassification rate for strong sensitizers and other sensitizers decreases, but the overclassification rate of nonsensitizers and weak sensitizers increases.

Figure 4. Overall Classification Rates for the LLNA EC3 Prediction of Human Potency for 112 Substances

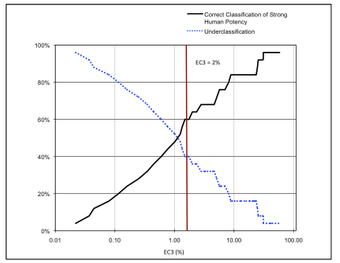


Abbreviations: EC3 = estimated substance concentration that produces a stimulation index of 3 in the murine local lymph node assay; LLNA = murine local lymph node assay.

Potency Prediction of Strong Sensitizers

- **Figure 5** shows the rates of correct classification and underclassification by the LLNA EC3 for the 25 strong human sensitizers.
  - 64% (16/25) of strong human sensitizers are also strong sensitizers in LLNA at EC3=2%
  - 36% (9/25) are under predicted by LLNA at EC3 = 2%

Figure 5. Classification Rates for LLNA EC3 Prediction of 25 Strong Human Sensitizers



Abbreviations: EC3 = estimated substance concentration that produces a stimulation index of 3 in the murine local lymph node assay; LLNA = murine local lymph node assay.

Prediction of Human Potency

- Classification rates for the LLNA EC3 values relative to strong and other human sensitizers and nonsensitizers are shown in **Table 2**.
  - Analysis of the complete database of 112 substances results in the following:
    - The optimum EC3 cutoff is 1.5% based on an overall correct classification rate of 62%.
    - The EC3 cutoff of 2% produced nearly the highest correct classification rate, 61%.
  - When the LLNA EC3 classification rates for the strong sensitizer, other sensitizer, and nonsensitizer categories are calculated separately:
    - The other sensitizer category is predicted better [77% (33/43) at EC3=2%] than the strong sensitizer category [64% (16/25) at EC3 = 2%].
    - Approximately one third of the strong human sensitizers are under classified as other sensitizers and nonsensitizers [36% (9/25) at EC3 = 2%].

Table 2. Classification Rates for LLNA EC3 Prediction of Human Potency for 112 Substances

EC3 Cutoff	Strong Human Sensitizers (threshold ≤500 μg/cm^2)[1]		Other Human Sensitizers (threshold >500 μg/cm^2)[1]			Human Nonsensitizers		Overall Correct Classification
	Correct	Under	Over	Correct	Under	Correct	Over	
Optimal cutoff EC3 = 1.5%	60% (15/25)	40% (10/25)	5% (2/43)	81% (35/43)	14% (6/43)	43% (19/44)	57% (25/44)	62% (69/112)
GHS cutoff EC3 = 2%	64% (16/25)	36% (9/25)	9% (4/43)	77% (33/43)	14% (6/43)	43% (19/44)	57% (25/44)	61% (68/112)

Abbreviations: EC3 = estimated substance concentration that produces a stimulation index of 3 in the murine local lymph node assay.

[1]Human induction concentration that produces a positive response in the human maximization test or human repeat insult patch test. Human induction threshold and LLNA EC3 values for the 61 sensitizers are shown in **Figure 3**.

Summary

- Prediction of Category 1A (strong) human sensitizers (n = 25) by LLNA EC3 ≤ 2%
  - 16 correct
  - 8 underclassified as other sensitizers (EC3 > 2%)
  - 1 misclassified by LLNA as a nonsensitizer
- Prediction of Category 1B (other) human sensitizers (n = 43) by LLNA EC3 > 2%
  - 33 correct
  - 4 overpredicted as strong sensitizers
  - 6 misclassified by LLNA as nonsensitizers
- Prediction of human nonsensitizers (n = 44) by LLNA
  - 19 correct
  - 25 false positives
    - 24 with EC3 >2%
    - 1 with EC3 ≤ 2%

Conclusions

- Over one-third of strong sensitizers would be underclassified as other skin sensitizers if the LLNA EC3 ≤ 2% were used to determine potency categories.
- The LLNA should not be considered as a stand-alone test to predict skin sensitization potency.
  - The LLNA EC3 ≤ 2% can be used as a screening assay to categorize a substance as a strong sensitizer.
  - However, EC3 > 2% should not be used to classify substances as other than strong sensitizers because it would result in over one third of strong sensitizers being underclassified as other sensitizers based on the available database.
- Other types of supporting information should be investigated for their usefulness in increasing the accuracy of categorization criteria for strong sensitizers.
  - For example, structure-activity relationships, peptide reactivity, human evidence, validated *in vitro* assays, historical data from related substances, other animal studies, etc.
- Information found to be useful should be incorporated into an integrated decision strategy for categorization.

References

Dean JH, Twerdok LE, Tice RR, Sailstad DM, Hattan DG, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay: II. Conclusions and recommendations of an independent scientific peer review panel. Regulatory Toxicology and Pharmacology 34(3): 258-273.

Haneke KE, Tice RR, Carson BL, Margolin BH, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay: III. Data analyses completed by the national toxicology program interagency center for the evaluation of alternative toxicological methods. Regulatory Toxicology and Pharmacology 34(3): 274-286.

ICCVAM. 1999. The Murine Local Lymph Node Assay: A Test Method for Assessing the Allergic Contact Dermatitis Potential of Chemicals/Compounds. The Results of an Independent Peer Review Evaluation Coordinated by ICCVAM and NICEATM (NIH 99-4494). Available at http://iccvam.niehs.nih.gov/docs/immunotox_docs/tlna/llnarep.pdf.

ICCVAM. 2008. Draft Background Review Document. Use of the Murine Local Lymph Node Assay (LLNA) to Determine Skin Sensitization Potency Categories. Available at http://iccvam.niehs.nih.gov/methods/immunotox/llna-panelDocs08.htm.

Kligman AM, Epstein W. 1975. Updating the maximization test for identifying contact allergens. Contact Dermatitis 1(4): 231-239.

NIEHS. 2007. The Murine Local Lymph Node Assay: Request for Comments, Nominations of Scientific Experts, and Submission of Data. *Federal Register* 72:27815-27817. Available at http://iccvam.niehs.nih.gov/SuppDocs/FedDocs/FR/FR_E7_9544.pdf.

Politano VT, Api AM. 2008. The Research Institute for Fragrance Materials' human repeated insult patch test protocol. Regulatory Toxicology and Pharmacology 52: 35-38.

Sailstad DM, Hattan D, Hill RN, Stokes WS. 2001. ICCVAM evaluation of the murine local lymph node assay: I. The ICCVAM review process. Regulatory Toxicology and Pharmacology 34(3): 249-257.

UN. 2009. Globally Harmonized System of Classification and Labelling of Chemicals (GHS). United Nations, Geneva, 559.

Acknowledgements

This poster was supported by the Intramural Research Program of the NIH, National Institute of Environmental Health Sciences. ILS staff were supported by NIEHS contract N01-ES 35504. The views expressed above do not necessarily represent the official positions of any U.S. Federal agency. Since this poster was written as part of the official duties of the authors, it can be freely copied.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
National Institutes of Health

NIEHS National Institute of Environmental Health Sciences

NTP National Toxicology Program